# A Nonlinear Generalization of Perron's theorem

A.Vladimirov V.I.Oseledets A.Rybko K.M.Khanin
D.V.Khmelev

26 June 2002

Published in "Doklady Mathematics", Vol.67, No. 2, 2003, p.221–225.

**Abstract**

A sufficient condition for global asymptotic stability of a mapping of a star-like surface in $\mathbb{R}^d_+$ into itself is given. This condition is based on the non-linear generalization of Perron's stability theorem.

Consider real vector space $\mathbb{R}^d$ with norm $\|x\| = |x_1| + \ldots + |x_d|$. A famous theorem by Perron [1] (see, e.g. [2]) gives some information on a matrix of a linear operator $A : \mathbb{R}^d_+ \to \mathbb{R}^d_+$, mapping the positive cone into itself.

**Theorem 1.** *Suppose that $A = (a_{ij}) > 0$ component-wise: $a_{ij} > 0$ for all $1 \leqslant i, j \leqslant d$. Then there exists a unique vector $x^* \in \mathbb{R}^d_+$, $x^* > 0$, $\|x\| = x^*_1 + \ldots + x^*_d = 1$, such that $x^*$ is an eigenvector for matrix $A$ and $x^*$ corresponds to a maximal eigenvalue, coinciding with spectral radius $\rho(A)$: $Ax^* = \rho(A)x^*$.*

*Moreover, for some $K > 0$, $0 < \beta < 1$ and for each $x \in \mathbb{R}^n_+ \setminus \{0\}$ the following inequality holds.*

$$\|A^n x/\|A^n x\| - x^*\| \leqslant K\beta^n. \tag{1}$$

There exists a classical generalization of this theorem, where the powers of $A$ in (1) are substituted with the product of an arbitrary sequence of matrices $A_i$ from the compact set $\mathcal{A}$ of component-wise positive matrices. In order to give the statement of the theorem the following definition is required.

Let us define *projective distance* $\rho(x, y)$ between non-zero $x$, $y \in \mathbb{R}^d \setminus \{0\}$ as follows

$$\rho(x, y) = \left\| \frac{x}{\|x\|} - \frac{y}{\|y\|} \right\|.$$

Notice that the projective distance depends on directions of $x$ and $y$, but not on their lengths. Hence projective distance is indeed the metric on the projective space $\mathbb{R}P^n_+$. For non-zero $x$ and $z$ we shall write $x \sim z$, if $x = az$ for some $a \in \mathbb{R}$.

**Theorem 2.** *Consider a family $\mathcal{A}$ of positive linear operators $A : \mathbb{R}^d_+ \to \mathbb{R}^d_+$ such that*

$$0 < B \leqslant A \leqslant D < \infty \text{ for all } A \in \mathcal{A}.$$

*Then there exist constants $\zeta = \zeta(B, D)$, $0 < \zeta < 1$, $N = N(B, D)$ such that for any $n > N$ and for arbitrary sequence of operators $A_1, \ldots, A_n \in \mathcal{A}$ and for any pair of non-zero vectors $x, y \geq 0$ the following inequality take place*

$$\rho(A_n \ldots A_1 x, A_n \ldots A_1 y) \leqslant \zeta \rho(x, y). \tag{2}$$

Theorem 2 is a "folk" theorem. In particular this means that we do not know the exact reference to the source where the theorem was proven. Therefore in the end of the paper we give the full proof for Theorem 2.

In this paper a non-linear generalization of Theorem 1 is given, arising naturally in the context of dynamical systems theory. To start with, consider $d$-dimensional cone $\mathbb{R}^d_+$ and fix a functional $\alpha : \mathbb{R}^d_+ \to \mathbb{R}_+$, satisfying the following conditions:

A1) $\alpha(\lambda x) = \lambda \alpha(x)$ for all $\lambda \geq 0$, $x \geq 0$

A2) $(\alpha'(x))^{\mathrm{T}} x > 0$ for all $x \geq 0$, $x \neq 0$ (where $^{\mathrm{T}}$ means transposition, and $\alpha'(x) = (\partial\alpha/\partial x_1, \ldots, \partial\alpha/\partial x_d)^{\mathrm{T}}$ is a derivative or a gradient of $\alpha$).

A3) $\alpha(x)$ is a smooth functional with third derivatives which are locally bounded.

Obviously $\alpha(x)$ induces a smooth star-like surface

$$S = \{x \in \mathbb{R}^d_+ \mid \alpha(x) = 1\}.$$

Consider the mapping $F : S \to \mathbb{R}^d_+$ and define the corresponding transformation $T : S \to S$:

$$Tx = \frac{F(x)}{\alpha(F(x))}.$$

It follows from A2) that transformation $T$ is the convolution of $F$ and the central projection to the surface $S$.

The normal to $S$ at point $x \in S$ is given by the derivative (the gradient) $\alpha'(x)$. Let us define

$$\gamma(x) = \alpha'(x)/((\alpha'(x))^{\mathrm{T}} x). \tag{3}$$

Notice that a normalizing factor in (3) is chosen in such a way that

$$\gamma(x)^{\mathrm{T}} x = 1.$$

**Theorem 3.** *Assume that*

$$A(x) = F'(x)(I - x\gamma(x)^{\mathrm{T}}) + F(x)\gamma(x)^{\mathrm{T}} \geq B > 0, \tag{4}$$

*where $I$ is the identity operator. Then there exist constants $N = N(F, \alpha)$ and $\lambda = \lambda(F, \alpha)$, $0 < \lambda < 1$, such that for any $n > N$ the following inequality*

$$\rho(T^n x, T^n y) \leqslant \lambda \rho(x, y) \tag{5}$$

*holds for all $x, y \in S$. It follows from* (5) *that transformation $T$ has a unique fixed point $x^* \in S$ and there exist constants $K > 0$, $0 < \beta < 1$ such that*

$$\|T^n x - x^*\| \leqslant K\beta^n \tag{6}$$

*for each $x \in S$.*

Notice that Theorem 3 is a non-linear generalization of Theorem 1. Indeed if $F(x) = Ax$ and $\alpha(x) = x_1 + \ldots + x_d$ then condition (4) holds with $B = A$, since $F'(x) = A = A(x)$.

**Remark 1.** *Notice that condition* (4) *depends on values of $F$ on the surface $S$ only. Indeed, $F'(x)$ acts in a tangent space at the point $x \in S$. But the operator $(I - x\gamma(x)^{\mathrm{T}})$ is a projection to the hyperplain $\gamma(x)^{\mathrm{T}} y = 0$ along the vector $x$ (since vector $x$ goes to zero).*

**Remark 2.** *Condition $B > 0$ in* (4) *can be weakened to inequality $B \geq 0$ provided that matrix $B$ is* primitive, *i.e., $B^l > 0$ for some $l > 0$. Further generalizations are possible.*

**Remark 3.** *Instead of the fixed mapping $F$ and the functional $\alpha$ one can consider a sequence $F_1$, $F_2$, ..., $\alpha_1$, $\alpha_2$, ..., which satisfies uniformly the conditions given. Then the corresponding sequence of transformations $T_i x = F_i(x)/\alpha_i(F_i(x))$ shall also be a contraction in $n$ steps:*

$$\rho(T_{i+n} \ldots T_{i+1} x, T_{i+n} \ldots T_{i+1} y) \leqslant \lambda \rho(x, y).$$

**Remark 4.** *Earlier some non-linear generalizations of Perron's theorem were given [3]. They are useful in certain problems of mathematical economics, mathematical demography and theory of difference equations [4]. However conditions on $F$ in papers [3, 4] are of global character and it's relation to differential condition* (4) *is not evident. One can verify that conditions* (4) *and [3] are equivalent for special case $d = 2$. At the same time the relationship between conditions* (4) *and [3] in general case is quite non-trivial. Indeed in order to verify their equivalence (or rather non-equivalence) in the case $d \geq 3$ one should compare two systems of inequalities of large dimension.*

**Remark 5.** *If the mapping $F$ is homogeneous of first degree, i.e., $F(\lambda x) = \lambda F(x)$ for all $\lambda \geq 0$, $x \geq 0$, then $F(x) = F'(x)x$ and condition* (4) *is reduced to the simple condition $F'(x) \geq B > 0$. In this case the statement about convergence to the fixed point (without estimate of the rate of convergence) was proven before (see [5]). Taking into account Remark 1 one can understand condition* (4) *as a statement of strict componentwise positivity of the derivative of homogeneous continuation for mapping $F$ from surface $S$. In principle this implies the convergence to a fixed point. However we believe that the proof suggested below has a number of obvious advantages. Firstly, it gives the estimate for the rate of convergence. Secondly, it can be applied for the sequence of different mappings. But most importantly this proof stress out the dynamical reasons for stability of the mapping $T$.*

**Remark 6.** *Inequality* (1) *in Theorem 1 follows from theorem 3. Indeed, since $A$ is a linear operator, inequality* (1) *is equivalent to inequality* (6), *if one takes $F(x) = Ax$ and $\alpha(x) = x_1 + \ldots + x_n$. Then $F'(x) = A$ and $A(x)$ in* (4) *is $A$. Therefore condition* (4) *holds with $B = A$.*

**Remark 7.** *If a mapping $F$ is first-degree homogeneous, i.e., $F(\lambda x) = \lambda F(x)$ for all $\lambda \geq 0$, $x \geq 0$, then $F(x) = F'(x)x$ and condition (4) is reduced to the simple inequality $F'(x) \geq B > 0$. In this case the convergence to a fixed point (without estimate of convergence rate) was proven before (see []). Remark 1 imply that condition (4) is a condition for component-wise positivity of degree-one homogenization of the mapping $F$ on the surface $S$. This observation, in principle, imply the convergence to a fixed point. However we believe that the proof suggested below takes a number of important advantages. First, it provides the estimate for the rate of convergence. Second, it can be used for a sequence of different mappings. And most important this proof highlights the dynamical nature for stability of the mapping $T$.*

*Proof of Theorem 3.* It follows from conditions of the Theorem that there exists a constant $C$ such that

$$\|T'(x)\| \leqslant C \text{ for each} x \in S \tag{7}$$

and

$$\frac{1}{C}\|x-y\| \leqslant \rho(x,y) \leqslant C\|y-x\| \text{ for all } x,y \in S. \tag{8}$$

Since $S$ is a smooth and compact surface, one can choose constant $C$ in such a way that inequalities (7) and (8) hold together with the following one

$$|\gamma(x)^{\mathrm{T}}(x-y)| \leqslant C\|x-y\|^2 \text{ for all } x,y \in S. \tag{9}$$

Inequality (9) can be interpreted geometrically as a condition of almost orthogonality of the normal $\gamma(x)$ and the vector $x-y$, such that $x-y$ is close to the tangent space at point $x$ to the surface $S$. Notice that

$$\begin{aligned} A(x)x &= (F'(x)(I - x\gamma(x)^{\mathrm{T}}) + F(x)\gamma(x)^{\mathrm{T}})x = \\ &= F'(x)(x - x\gamma(x)^{\mathrm{T}}x) + F(x)\gamma(x)x = F(x). \end{aligned}$$

Hence

$$Tx \sim A(x)x = F(x),$$

where the equivalence relation $\sim$ was defined above. Consider trajectories $x_i = T^i x_0$, $y_i = T^i y_0$, where $x_0 = x$ and $y_0 = y$. Then

$$\begin{aligned} x_i &\sim A(x_{i-1})\ldots A(x_0)x_0, \\ y_i &\sim A(y_{i-1})\ldots A(y_0)y_0. \end{aligned}$$

Denote

$$A_i^j(x_j) = A(x_{j-1})\ldots A(x_i),\ A_j^j = I.$$

Let us estimate the projective distance between $A_0^n(x_0)y_0$ and $T^n y_0$:

$$\begin{aligned}\rho(T^n y_0, A_0^n(x_0)y_0) &\leqslant \sum_{i=1}^{n} \rho(A_i^n(x_i)T^i y_0, A_{i-1}^n(x_{i-1})T^{i-1}y_0) = \\ &= \sum_{i=1}^{n} \rho(A_i^n(x_i)TT^{i-1}y_0, A_i^n(x_i)A(x_{i-1})T^{i-1}y_0) \leqslant \\ &\leqslant \sum_{i=1}^{n} K_1^{n-i}\rho(Ty_{i-1}, A(x_{i-1})y_{i-1}),\end{aligned}$$

where $K_1 = K_1(C)$. Since $Ty_i \sim A(y_{i-1})y_{i-1}$, we obtain

$$\begin{aligned}\rho(Ty_{i-1}, A(x_{i-1})y_{i-1}) = \rho(A(y_{i-1})y_{i-1}, A(x_{i-1})y_{i-1}) \leqslant \\ \leqslant C\|A(y_{i-1})y_{i-1} - A(x_{i-1})y_{i-1}\|.\end{aligned}$$

Notice that

$$\begin{aligned}&A(v)v - A(u)v = \\ &\quad = F(v) - (F'(u)(I - u\gamma(u)^{\mathrm{T}}) + F(u)\gamma(u)^{\mathrm{T}})((v-u)+u) = \\ &\quad = F(v) - F(u) - (F'(u) + (F(u) - F'(u)u)\gamma(u)^{\mathrm{T}})(v-u) = \\ &\quad = F(v) - F(u) - F'(u)(v-u) - (F(u) - F'(u)u)\gamma(u)^{\mathrm{T}}(v-u).\end{aligned}$$

Taking $u = x_{i-1}$, $v = y_{i-1}$ and using the smoothness of $F$ together with estimate (9), we obtain

$$\rho(Ty_{i-1}, A(x_{i-1})y_{i-1}) \leqslant C_1\|x_{i-1} - y_{i-1}\|^2.$$

Therefore

$$\begin{aligned}\rho(T^n y_0, A_0^n y_0) &\leqslant \sum_{i=1}^{n} K_1^{n-i}C_1\|T^{i-1}x_0 - T^{i-1}y_0\|^2 \leqslant \\ &\leqslant C_1\sum_{i=1}^{n} K_1^{n-i}C^{2i}\|x_0 - y_0\|^2 \leqslant C_2(n)\|x_0 - y_0\|^2 \leqslant \\ &\leqslant C_3(n)(\rho(x,y))^2.\end{aligned}$$

It remains to notice that

$$\rho(T^n x_0, T^n y_0) \leqslant \rho(T^n y_0, A_0^n(x_0)y_0) + \rho(A_0^n(x_0)y_0, T^n x_0).$$

Since $T^n x_0 \sim A_0^n(x_0)x_0$, using Theorem 2, we obtain

$$\rho(A_0^n(x_0)y_0, T^n x_0) = \rho(A_0^n(x_0)y_0, A_0^n(x_0)x_0) \leqslant \zeta\rho(x_0, y_0).$$

Hence

$$\rho(T^n x_0, T^n y_0) \leqslant \zeta\rho(x_0, y_0) + C_3(n)\rho(x_0, y_0)^2 = \\ = (\zeta + C_3(n)\rho(x_0, y_0))\rho(x_0, y_0)$$

Choose $\varepsilon > 0$ such that

$$\lambda = \zeta + C_3(n)\rho(x_0, y_0) < 1.$$

Then $\rho(T^n x_0, T^n y_0) < \lambda\rho(x_0, y_0)$ if $\rho(x_0, y_0) < \varepsilon$. Obviously the local contraction obtained together with compactness of $S$ implies the global contraction (5).

Since trajectories of any two points under iterations by $T^n$ converge exponentially, the transformation $T$ has a unique fixed point $x^*$, which is globally asymptotically stable. □

Theorem 3 implies the following corollary, which can be applied in economical and demographic problems [4]. More precisely, suppose that $F(x)$ is a *concave* mapping of $\mathbb{R}^d_+$ into itself:

$$F(\lambda x + (1-\lambda)y) \geq \lambda F(x) + (1-\lambda)F(y), \ 0 \leqslant \lambda \leqslant 1, x, y \in \mathbb{R}^d_+. \tag{10}$$

Assume also that $F$ also is component-wise monotone with strictly positive derivative

$$F(x) \leqslant F(y) \text{ for all } x \leqslant y, x, y \in \mathbb{R}^d_+ \text{ and } F'(x) > 0 \text{ for all } x \in \mathbb{R}^d_+, \tag{11}$$

and, moreover, assume that $\alpha(x)$ satisfy conditions (A1)–(A3) and $\alpha'(x) \geq 0$ (which also means component-wise monotonicity).

**Theorem 4.** *For arbitrary $\alpha$ as assumed and for arbitrary $F$ such that conditions* (10) *and* (11) *hold, the condition* (4) *of Theorem 3 is fulfilled.*

*Proof.* Let $y = 0$ in condition (10):

$$F(\lambda x) \geq \lambda F(x). \tag{12}$$

Notice that in first order

$$F(\lambda x) = F(x) + F'(x)(\lambda x - x) + O(\|\lambda x - x\|^2).$$

Substituting this expression into (12), we obtain

$$F(x) + (\lambda - 1)F'(x)x + O((\lambda - 1)^2\|x\|^2) \geq \lambda F(x).$$

Therefore

$$(1-\lambda)F(x) + O((1-\lambda)^2\|x\|^2) \geq (1-\lambda)F'(x)x.$$

Dividing over $1 - \lambda$ yields

$$F(x) + O((1-\lambda)\|x\|^2) \geq F'(x)x.$$

Taking the limit $\lambda \to 1$, we obtain

$$F(x) \geq F'(x)x.$$

Notice now that condition (4) can be given in the following form

$$A(x) = F'(x) + (F(x) - F'(x)x)\gamma(x)^{\mathrm{T}} \geq B > 0.$$

Obviously the inequalities given above, (11) and compactness of $S$ imply the last estimate. $\square$

In conclusion we give the proof for Theorem 2. Essentially theorem 2 follows naturally from results presented by E.Seneta [6, Section 3.2] in the proof of the contracting property for the products of non-negative matrices. We start with the following lemma, which goes back to Markov [7].

**Lemma 1 (Markov, [7, 6]).** *Let $P = (p_{ij})_{i,j=1}^{d}$ be a stochastic matrix, $\sum_{j=1}^{d} p_{ij} = 1$, and $w \in \mathbb{R}^d$ a $d$-component vector (not necessarily non-negative), having maximum component $M_0$ and minimum component $m_0$; and let $M_1$ and $m_1$ be the maximum and minimum components of the vector $Pw$. Then $M_1 \leqslant M_0$, $m_1 \geq m_0$ and*

$$M_1 - m_1 \leqslant (1 - 2\varepsilon)(M_0 - m_0),$$

*where $\varepsilon \geq 0$ is the minimum entry of $P$.*

Let $a_{i,j}^{(m)}$ be an element of $A_m$ in the row $i$ and column $j$. Denote

$$A^{(k,l)} = A_l \dots A_{k+1}, \; k < l$$

for matrix product. Elements of the matrix $A^{(k,l)}$ are denoted by $a_{i,j}^{(k,l)}$:

$$A^{(k,l)} = (a_{i,j}^{(k,l)})_{i,j=1}^{d}.$$

The next lemma is given without proof since its statement follows easily from results proven in [6, 8].

**Lemma 2.** *There exist constants $K_1 = K_1(B, D) > 0$, $K_2 = K_2(B, D) > 0$, $\lambda = \lambda(B, D) > 0$ such that for arbitrary product of matrices, satisfying conditions of theorem 2 the following statements hold.*

*1)* $\dfrac{1}{K_1} \leqslant \dfrac{a_{s,i}^{(k,l)}}{a_{s,j}^{(k,l)}}, \dfrac{a_{i,s}^{(k,l)}}{a_{j,s}^{(k,l)}} \leqslant K_1$ *for all* $1 \leqslant s, i, j \leqslant d$.

*2) For all* $1 \leqslant s, r, i, j \leqslant d$

$$\left| \frac{a_{s,i}^{(k,l)}}{a_{s,j}^{(k,l)}} - \frac{a_{r,i}^{(k,l)}}{a_{r,j}^{(k,l)}} \right| \leqslant K_2 \lambda^{k-l-1}$$

**Lemma 3.** *The following decomposition takes place*

$$A^{(0,n)} = a_{1,1}^{(0,n)}(u^{(n)}(w^{(n)})^{\mathrm{T}} + \Omega^{(n)}),$$

*where*

$$\frac{1}{K_1} \leqslant u_i^{(n)}, w_i^{(n)} \leqslant K_1,$$

$$\Omega_{ij}^{(n)} \leqslant K_1 K_2 \lambda^n.$$

*Proof.* Take

$$u^{(n)} = \begin{pmatrix} a_{11}^{(0,n)}/a_{11}^{(0,n)} \\ \vdots \\ a_{d1}^{(0,n)}/a_{11}^{(0,n)} \end{pmatrix}, \ (w^{(n)})^{\mathrm{T}} = (a_{11}^{(0,n)}/a_{11}^{(0,n)}, \dots, a_{1d}^{(0,n)}/a_{11}^{(0,n)}),$$

i.e.,

$$u_i^{(n)} = \frac{a_{i,1}^{(0,n)}}{a_{1,1}^{(0,n)}}, \ w_i^{(n)} = \frac{a_{1,i}^{(0,n)}}{a_{1,1}^{(0,n)}}.$$

Then

$$\Omega_{ij}^{(n)} = \left(\frac{1}{a_{1,1}^{(0,n)}} A^{(0,n)} - u^{(n)}(w^{(n)})^{\mathrm{T}}\right)_{ij} = \frac{a_{i,j}^{(0,n)}}{a_{1,1}^{(0,n)}} - \frac{a_{i,1}^{(0,n)}}{a_{1,1}^{(0,n)}} \frac{a_{1,j}^{(0,n)}}{a_{1,1}^{(0,n)}}.$$

It follows from statements 1), 2) of Lemma 2 that the following inequality holds

$$|\Omega_{ij}^{(n)}| = \frac{a_{i,1}^{(0,n)}}{a_{1,1}^{(0,n)}} \cdot \left| \frac{a_{i,j}^{(0,n)}}{a_{i,1}^{(0,n)}} - \frac{a_{1,j}^{(0,n)}}{a_{1,1}^{(0,n)}} \right| \leqslant K_1 K_2 \lambda^{n-1}. \qquad \square$$

*Proof of the theorem 2.* Without loss of generality we shall assume that $\|x\| = \|y\| = 1$. Then $\rho(x, y) = \|x - y\|$. By Lemma 3 the following decomposition takes place

$$A^{(0,n)} = a(\Omega + uw^{\mathrm{T}}),$$

where $1/K_1 \leqslant u_i, w_i \leqslant K_1$, $\|\Omega\| \leqslant K_3 \lambda^n$.

$$\begin{aligned} \rho(A^{(0,n)}x, A^{(0,n)}y) &= \rho((\Omega + uw^{\mathrm{T}})\frac{x}{w^{\mathrm{T}}x}, (\Omega + uw^{\mathrm{T}})\frac{y}{w^{\mathrm{T}}y}) = \\ &= \rho(\Omega\frac{x}{w^{\mathrm{T}}x} + u, \Omega\frac{y}{w^{\mathrm{T}}y} + u) = \\ &= \|G(H(w(y))) - G(H(w(x)))\|, \end{aligned}$$

where $G(z) = z/\|z\|$, $H(x) = \Omega x + u$, $w(x) = x/w^{\mathrm{T}}x$. Newton-Leibniz formula says

$$G(H(w(y))) - G(H(w(x))) = \int_0^1 G'(H(w(\gamma(s))))H'(w(\gamma(s)))w'(\gamma(s))\gamma'(s)ds,$$

where $\gamma(s) = (1-s)x + sy$. Notice that

$$\begin{aligned} \gamma'(s) &= y - x \text{ and } \|\gamma'(s)\| = \|y-x\|, \\ w'(\gamma(s)) &\leqslant \overline{w}(B,D), \\ H'(w(\gamma(s))) &= \Omega \text{ and } \|\Omega\| \leqslant K_3\lambda^n, \\ G'(H(w(\gamma(s)))) &= G'(Y(\gamma(x))) \text{ and } \|G'\| \leqslant \overline{G}(B,D), \end{aligned}$$

where second and fourth bounds for derivatives follow from the following facts: $\gamma(s) \in \Delta = \{x \geq 0 | x_1 + \ldots + x_d = 1\}$; vector $w$ is uniformly bounded from zero; and the mapping $x \mapsto H(w(x))$ is the mapping $x \mapsto A^{(0,n)}x/(a_{1,1}^{(0,n)} w^{\mathrm{T}} x)$, which is uniformly positive and uniformly bounded by Lemma 2. Hence

$$\begin{aligned} \|G(H(w(y))) - G(H(w(x)))\| &\leqslant \overline{G}(B,D)\|\Omega\|\overline{w}(B,D)\|y-x\| \leqslant \\ &\leqslant K_4(B,D)\lambda^n\|y-x\| = K_4(B,D)\lambda^n\rho(x,y). \qquad \square \end{aligned}$$

# References


[1] O. Perron. Zur Theorie der Matrizes. *Math. Ann. 64, 248-263*, (1907).

[2] R.A Horn and C.R. Johnson. *Matrix Analysis.* Cambridge Univ. Press, Cambridge, 1989.

[3] Ulrich Krause. Perron's stability theorem for nonlinear mappings. *J. Math. Econom.*, 15(3):275–282, 1986.

[4] Ulrich Krause. Concave Perron-Frobenius Theory and Applications. *Nonlinear analysis*, 47:1457–1466, 2001.

[5] Hukukane Nikaidô. *Convex structures and economic theory.* Mathematics in Science and Engineering, Vol. 51. Academic Press, New York, 1968.

[6] E. Seneta. *Non-negative matrices.* Halsted Press [A division of John Wiley & Sons], New York, 1973. An introduction to theory and applications.

[7] A.A. Markov. Investigation of an important case of dependent trials. *Izvestia Akad. Nauk. series 6*, (1):61–80, 1907.

[8] D. M. Hardcastle and K. Khanin. Continued fractions and the $d$-dimensional Gauss transformation. *Comm. Math. Phys.*, 215(3):487–515, 2001.